Decomposition of Overlapping Protein Complexes

A Graph Theoretical Method for Analyzing Static and Dynamic Protein Associations

Teresa Przytycka

NIH / NLM / NCBI

joint work with

Elena Zotenko
Katia S. Guimaraes
Raja Jothi

Investigating protein-protein interaction networks

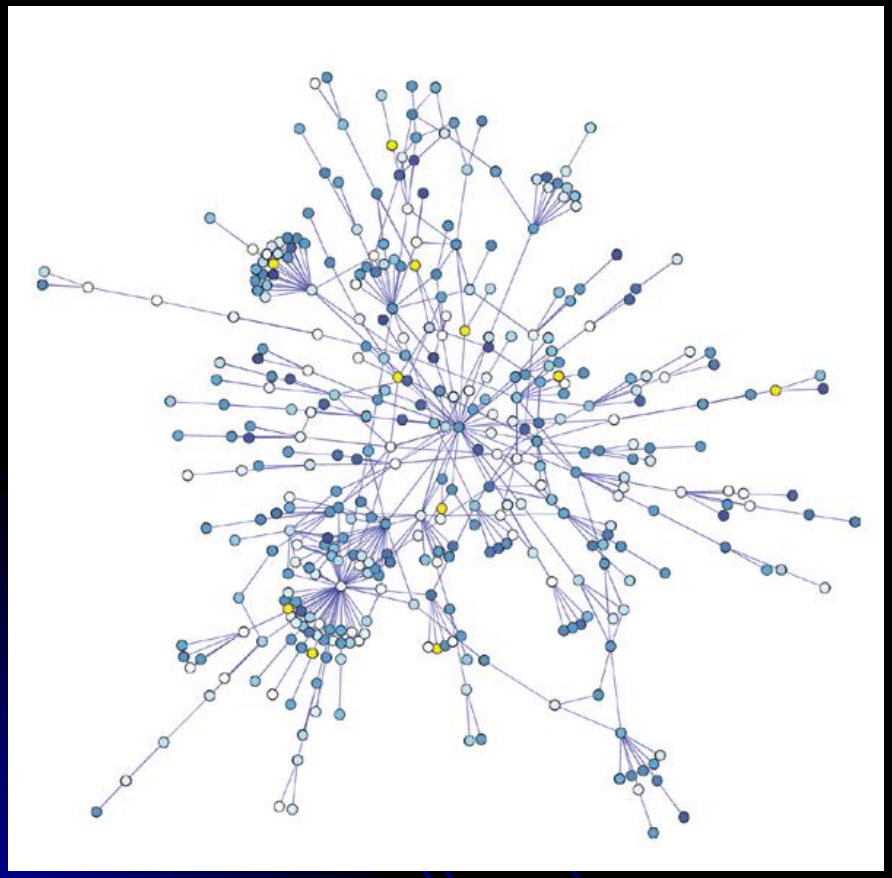


Image by Gary Bader (Memorial Sloan-Kettering Cancer Center).

- Scale free vertex degree distribution

  (Barabasi & Albert 1999; …)
- Modular organization
- Recognizing functional modules

  (Spirin & Mirny 2003, Rives & Galitski 2003, Bader *et al.* 2003, Bu *et al.* 2003,…)
- Representing variants of protein complexes

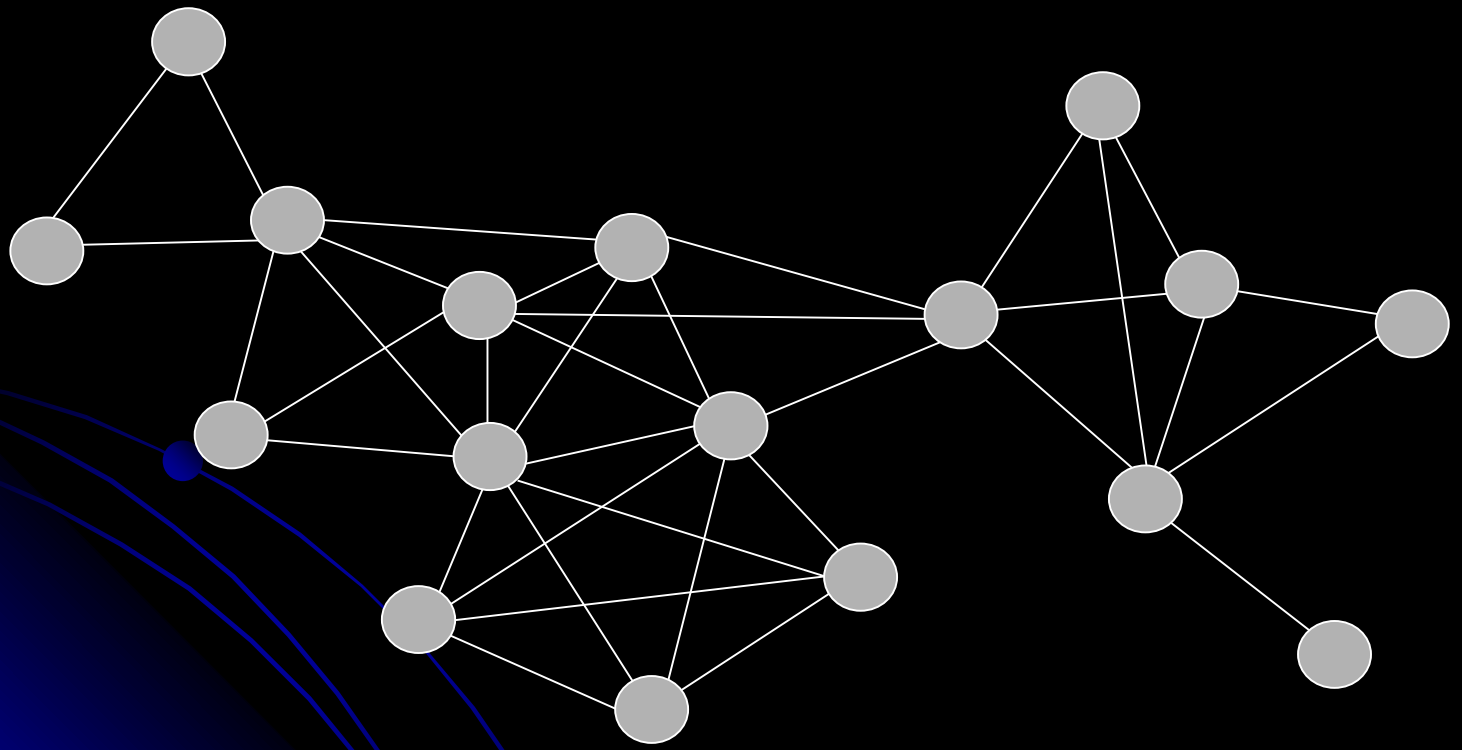
  (Gagnuer *et al., 2004)*

Functional Modules and Functional Groups

- Functional Module: Group of genes or their products in a metabolic or signaling pathway, which are related by one or more genetic or cellular interactions and whose members have more relations among themselves than with members of other modules (Tornow *et al.* 2003)
- Functional Group: protein complex (alternatively a group of pairwise interacting proteins) or a set of alternative variants of such a complex.
- Functional group is part of functional module

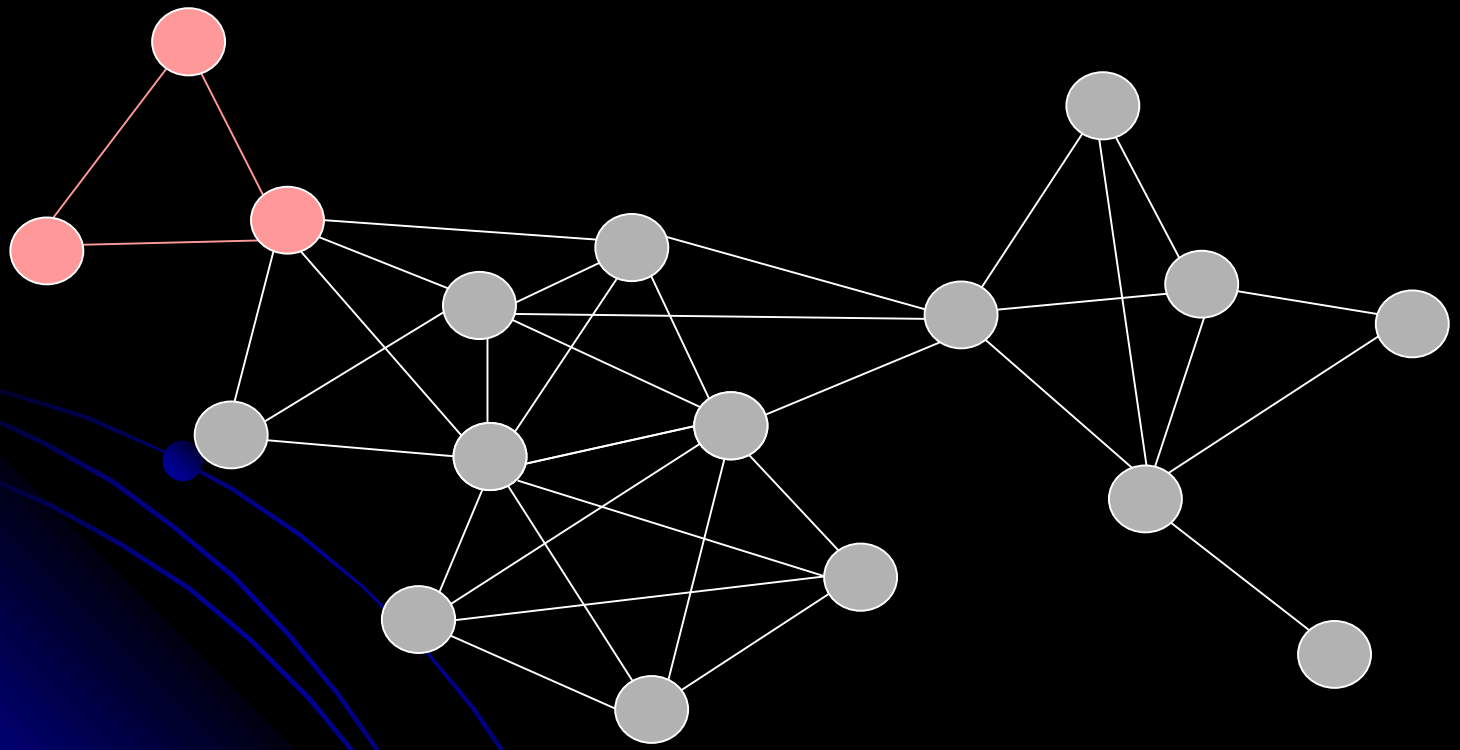
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques



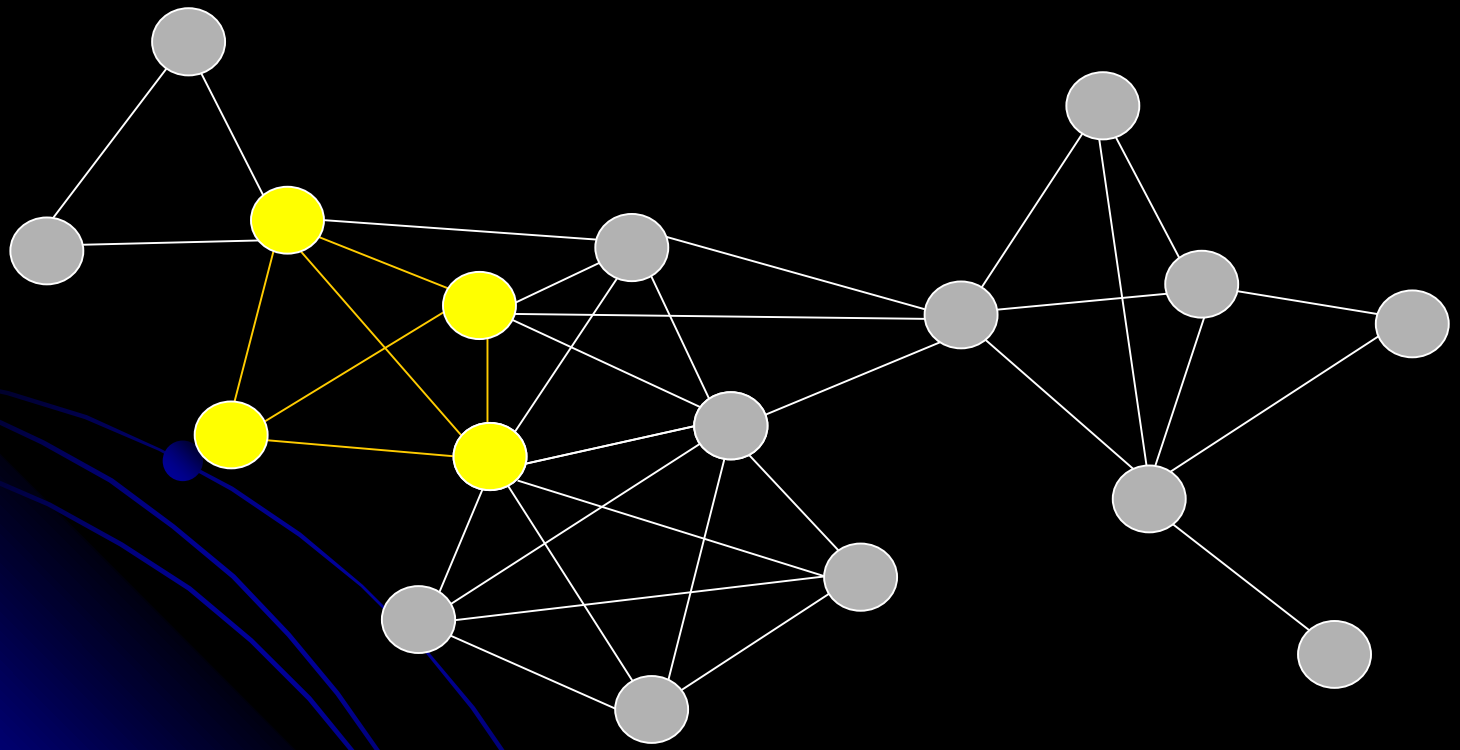
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques



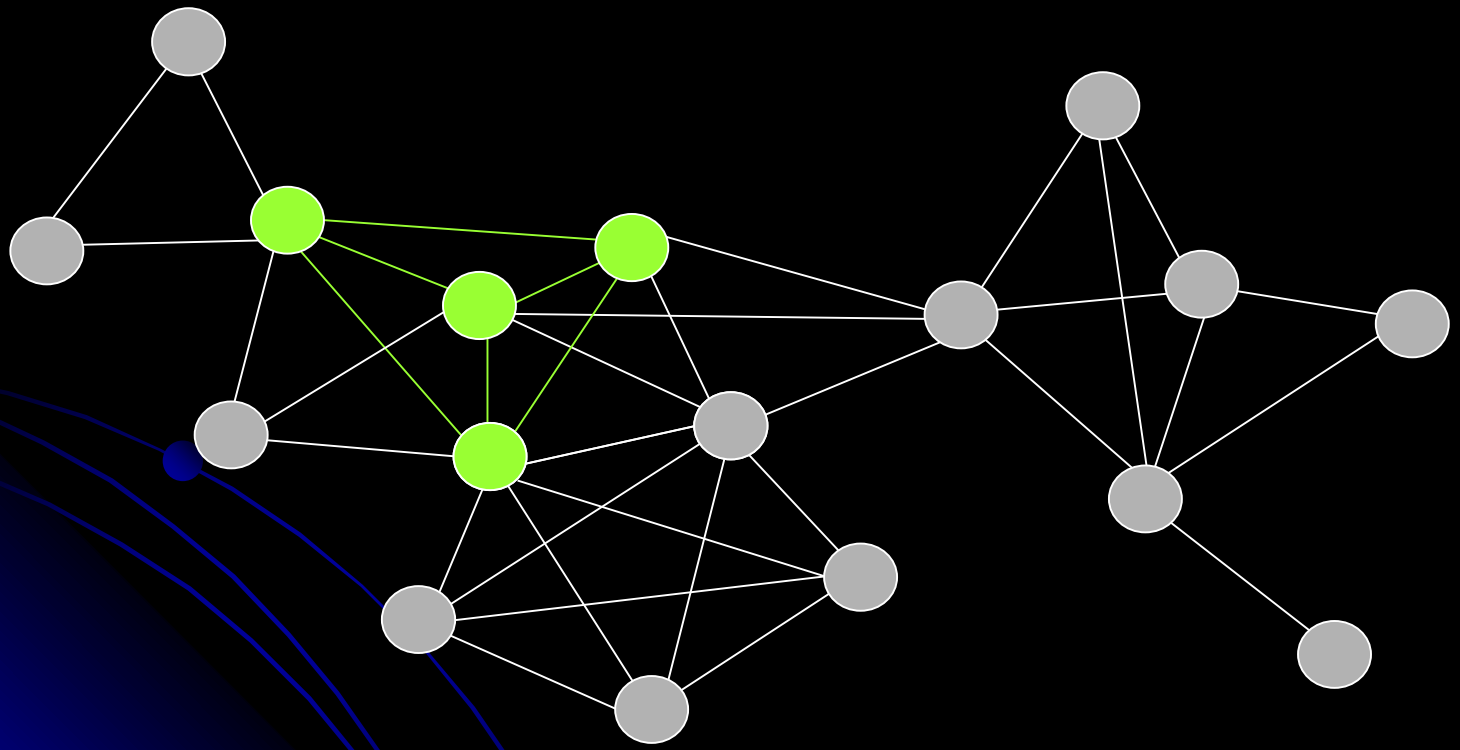
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques



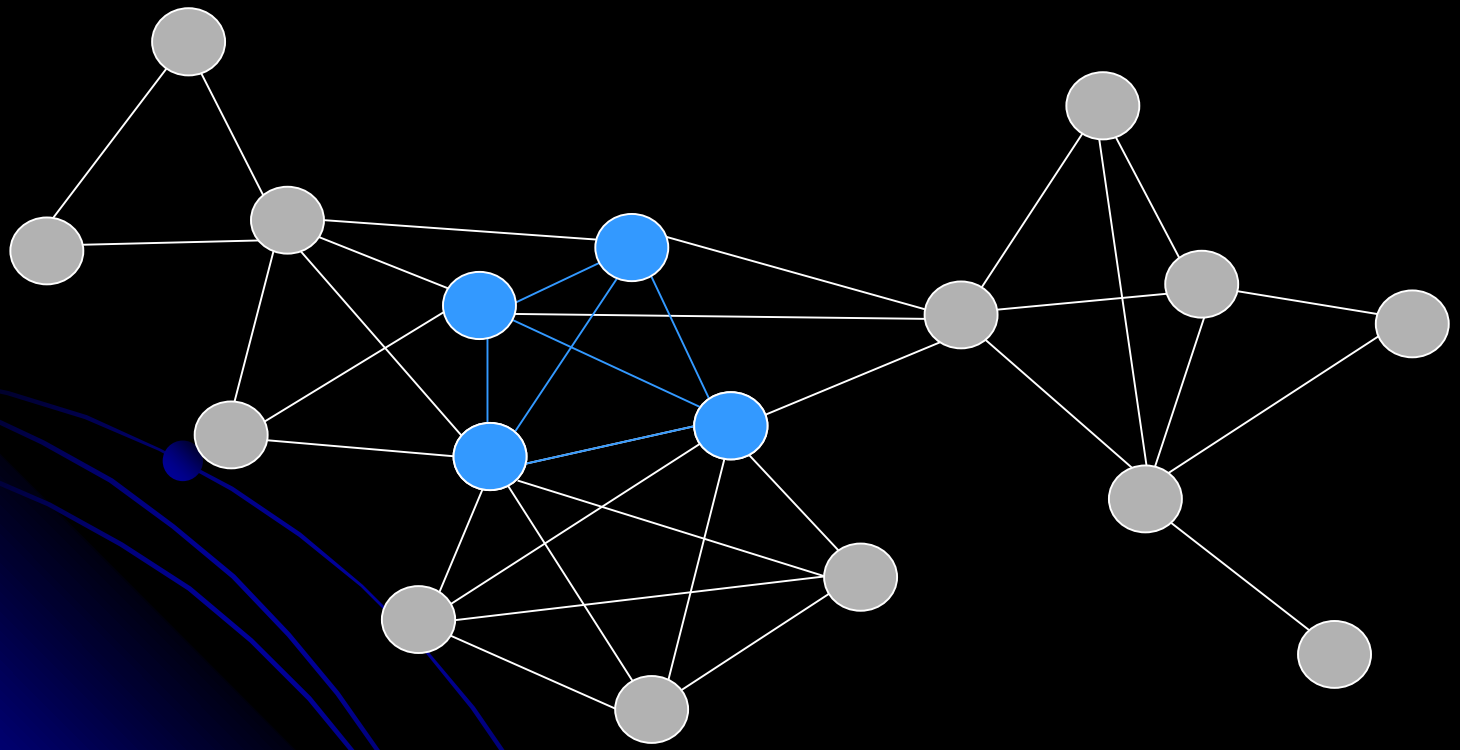
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques



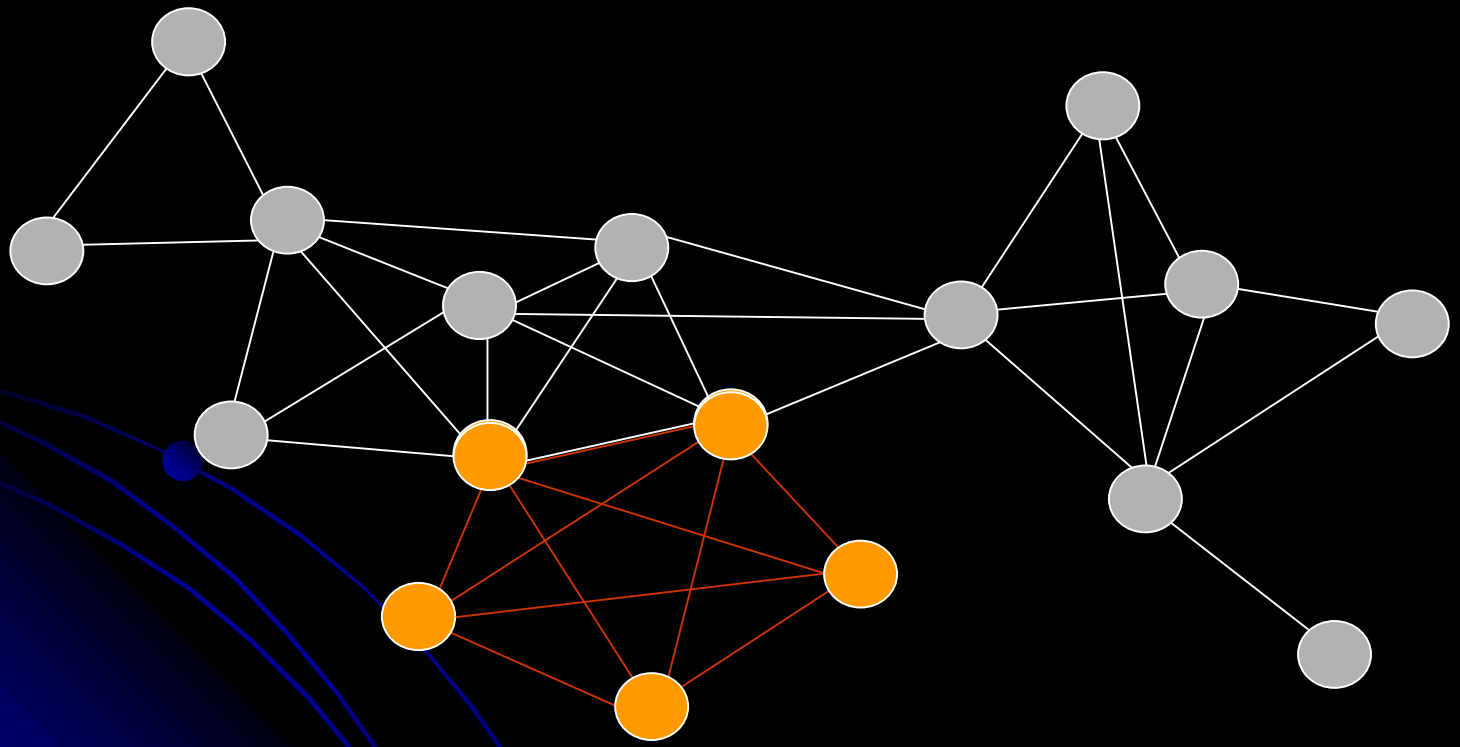
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques



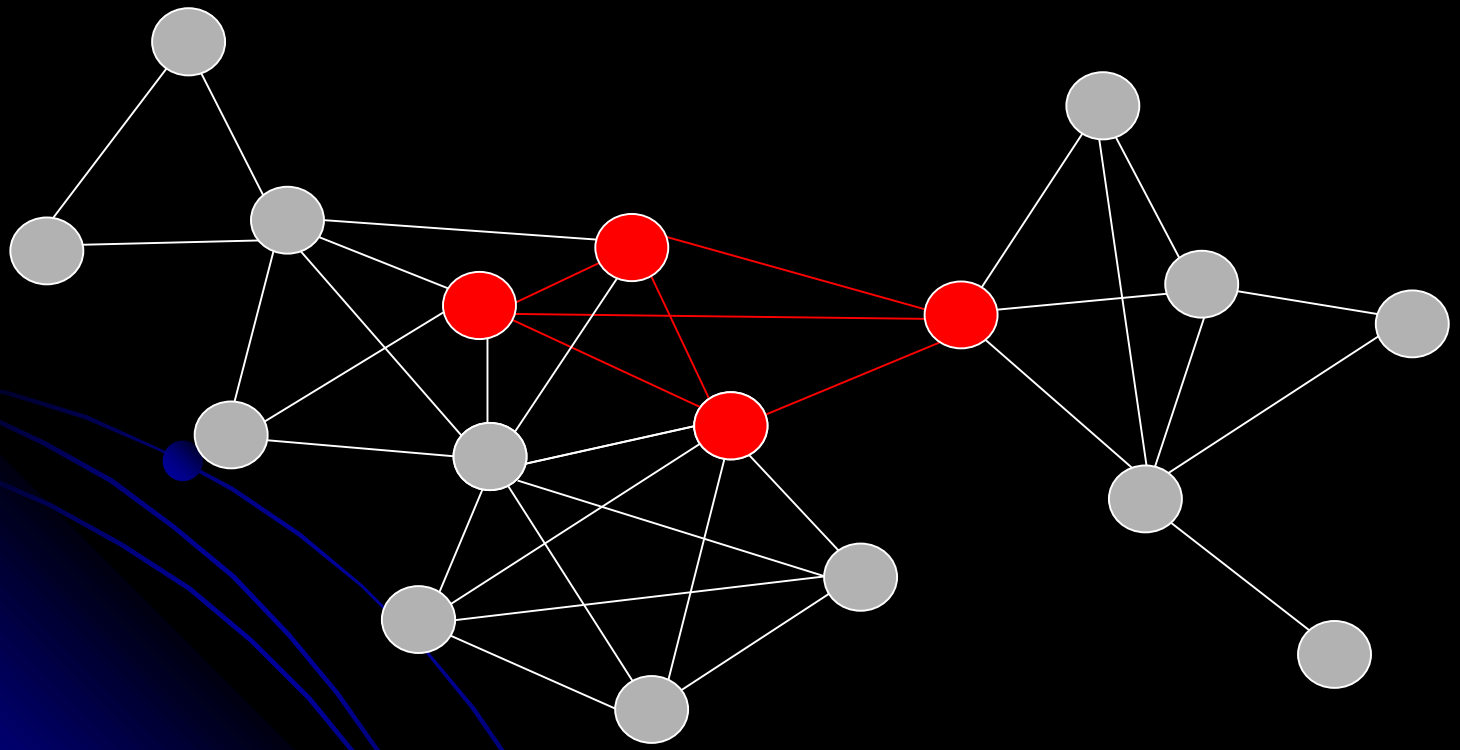
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques



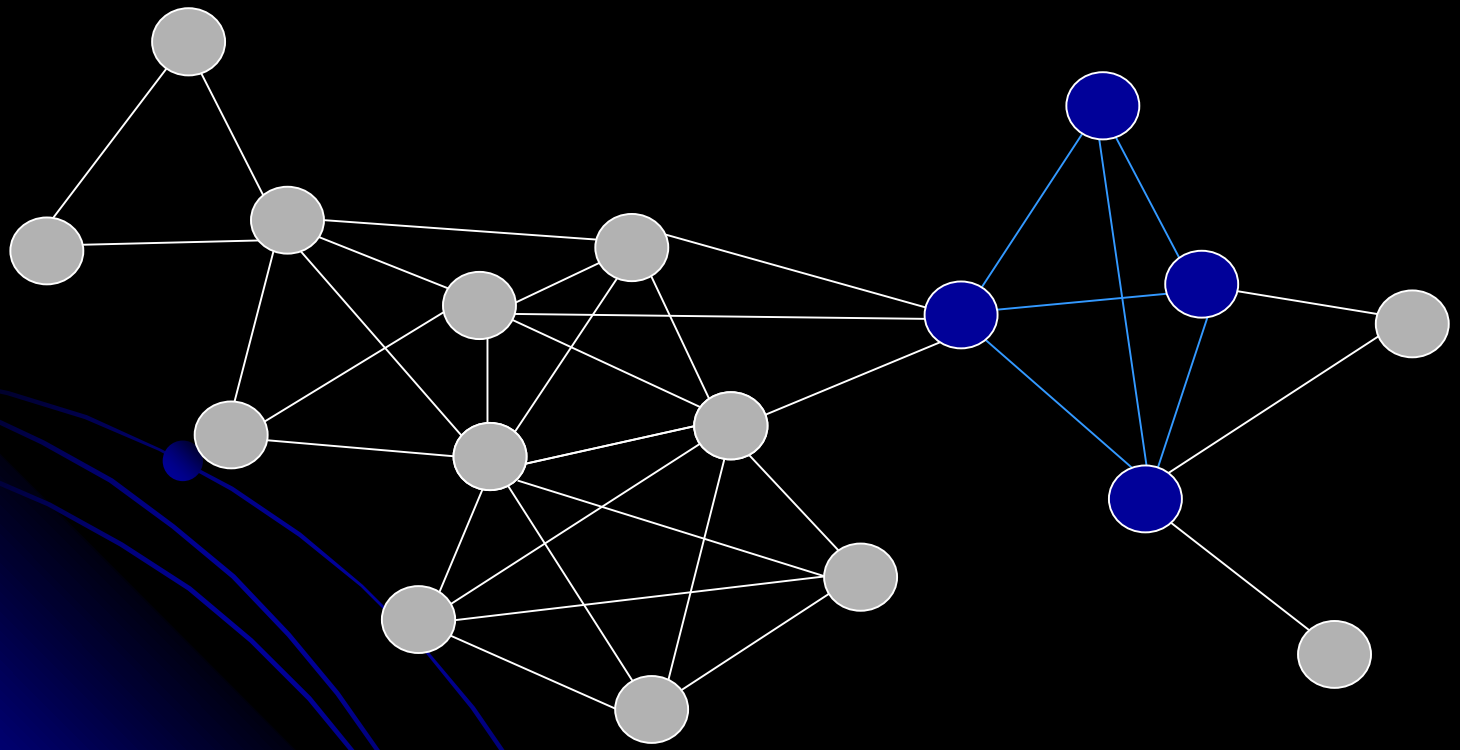
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques



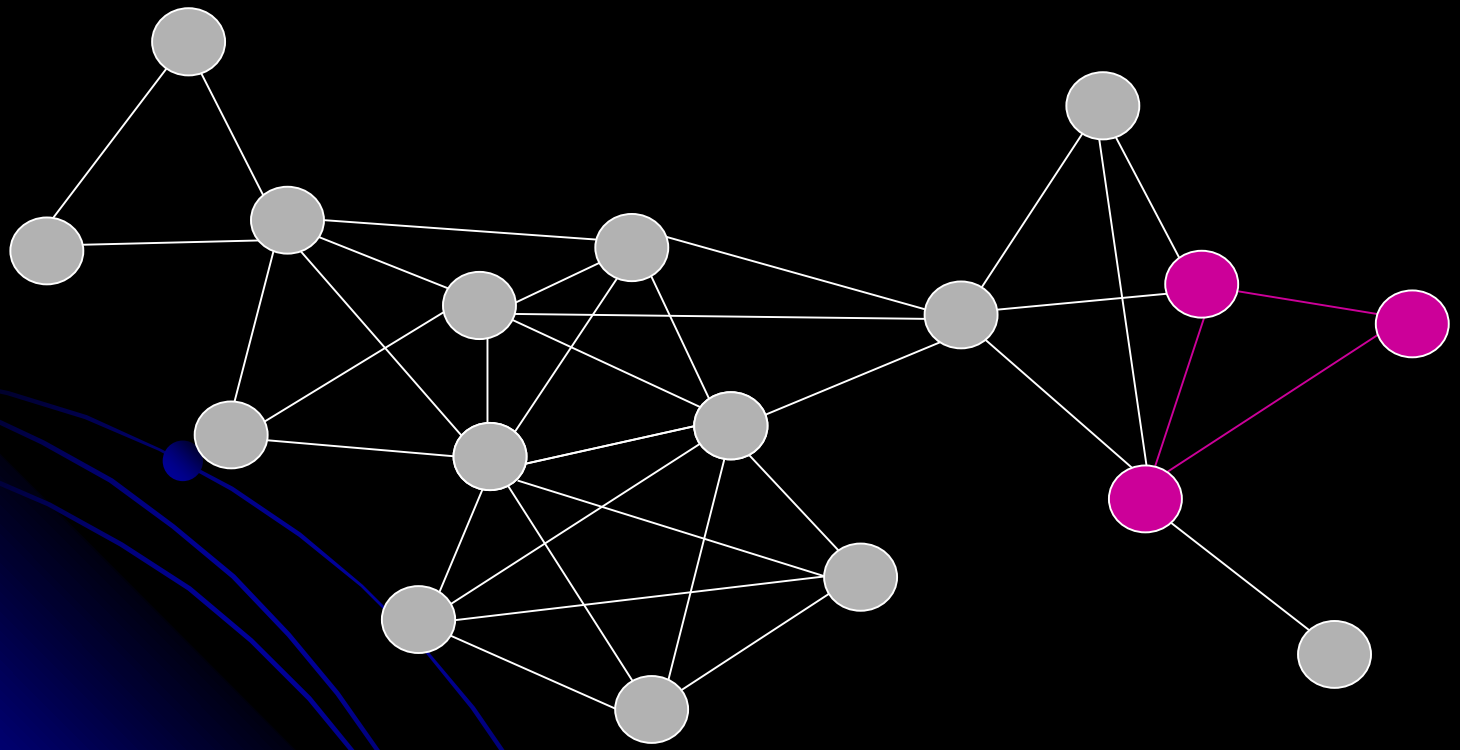
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques



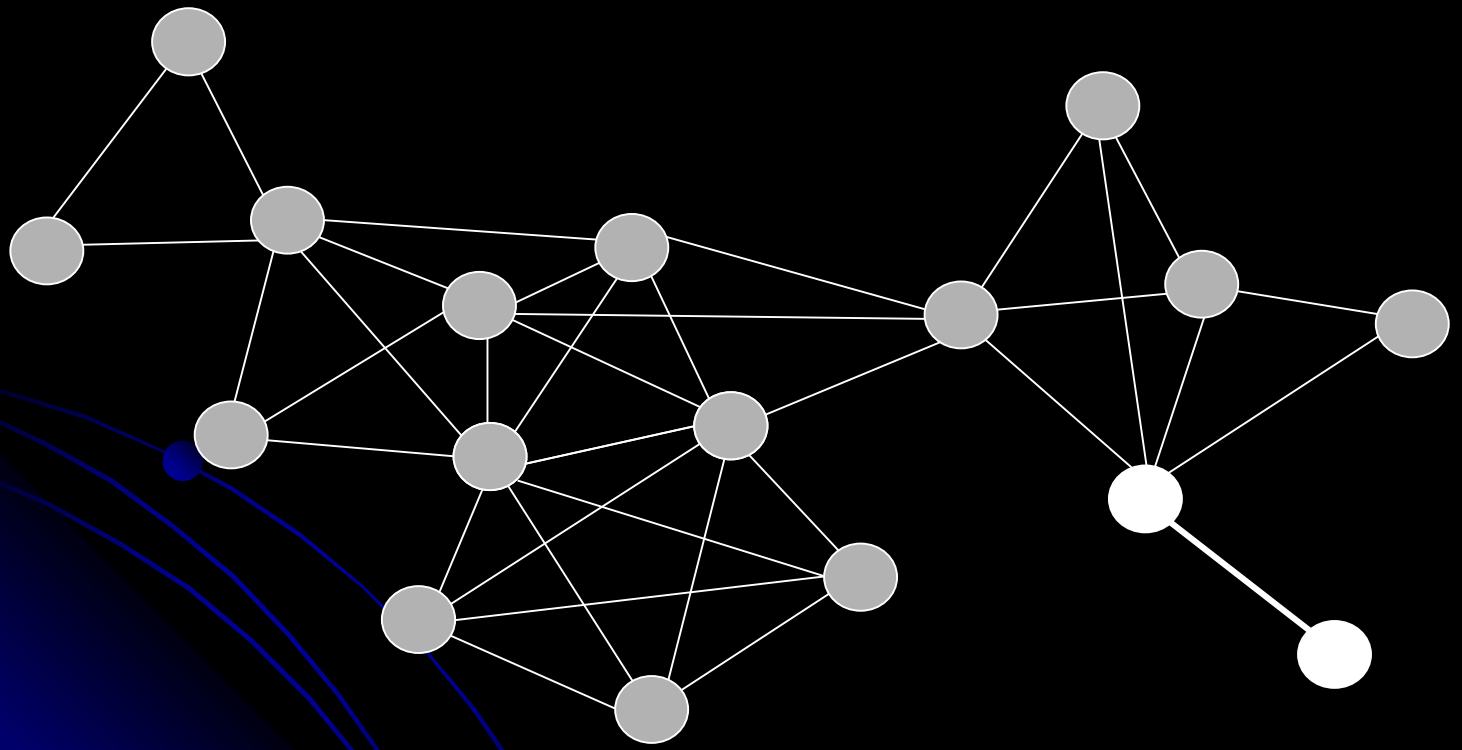
Overlaps between Functional Groups

For an illustration functional groups = maximal cliques

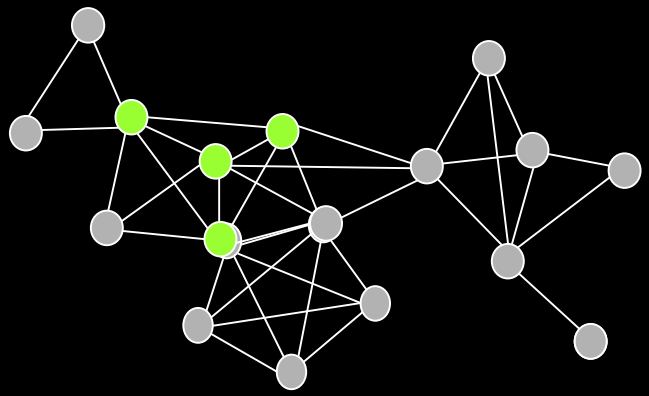


Overlaps between Functional Groups

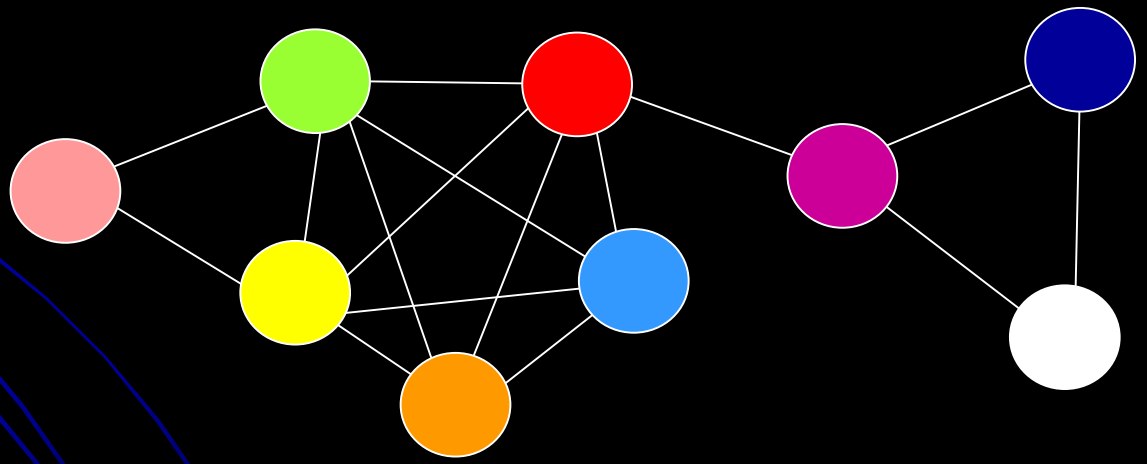
For an illustration functional groups = maximal cliques



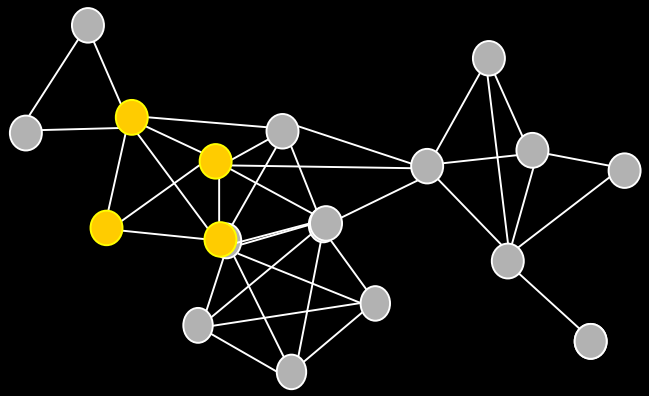
First line of attack



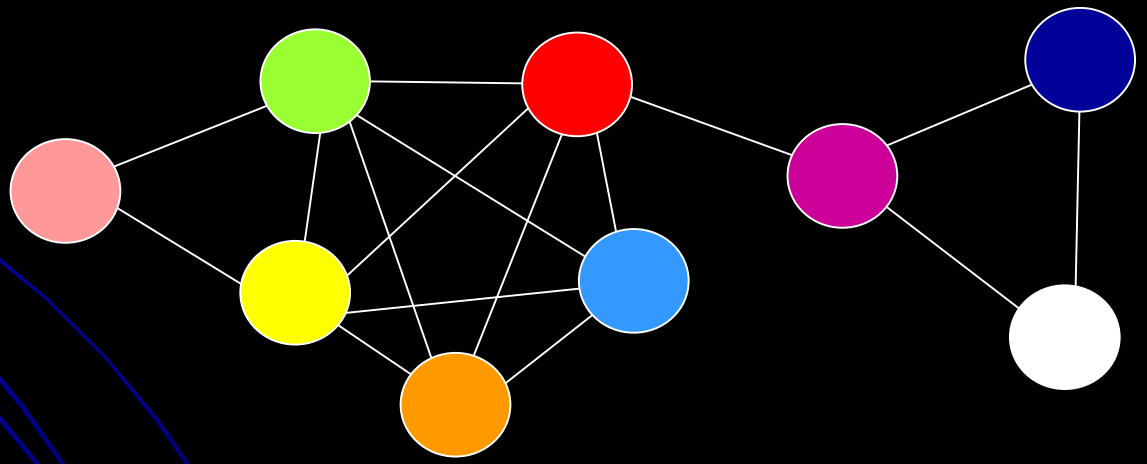
Overlap graph:
Nodes= functional groups
Edges= overlaps between them



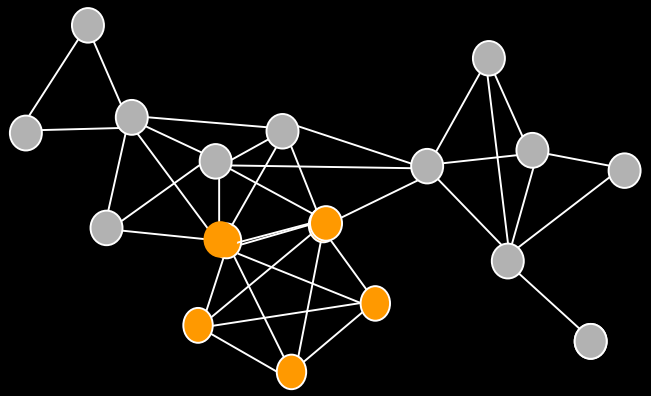
First line of attack



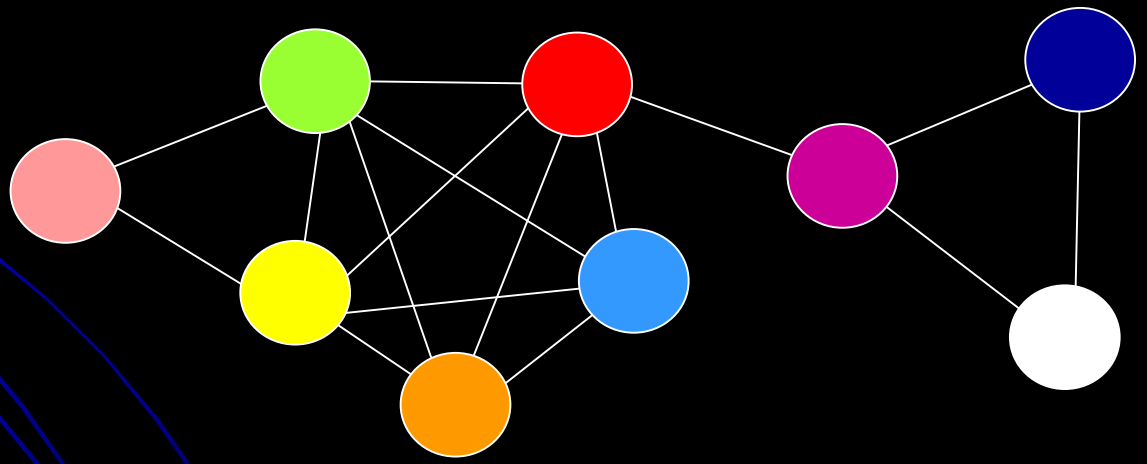
Overlap graph:
Nodes= functional groups
Edges= overlaps between them



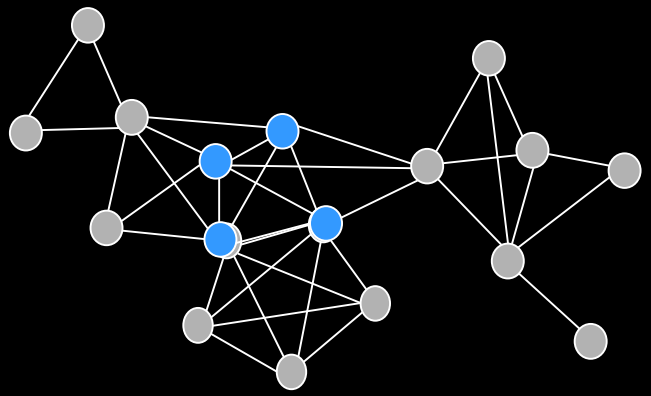
First line of attack



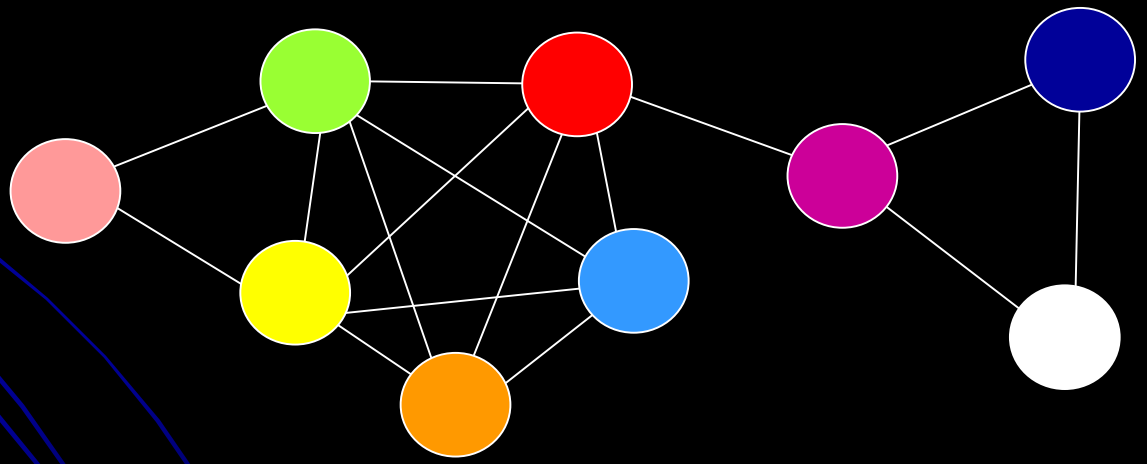
Overlap graph:
Nodes= functional groups
Edges= overlaps between them



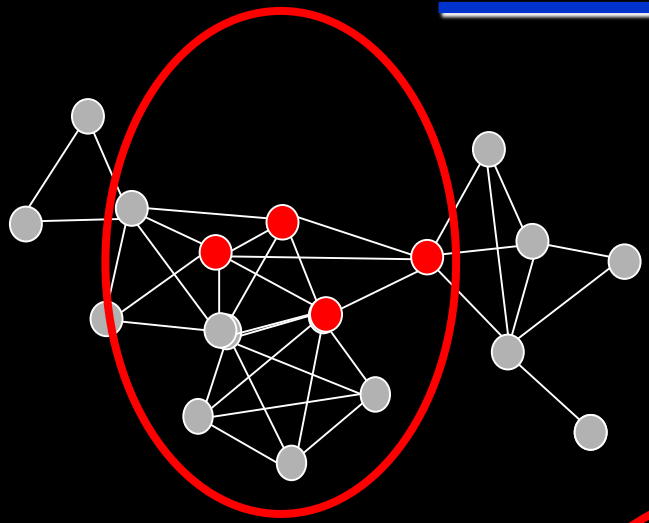
First line of attack



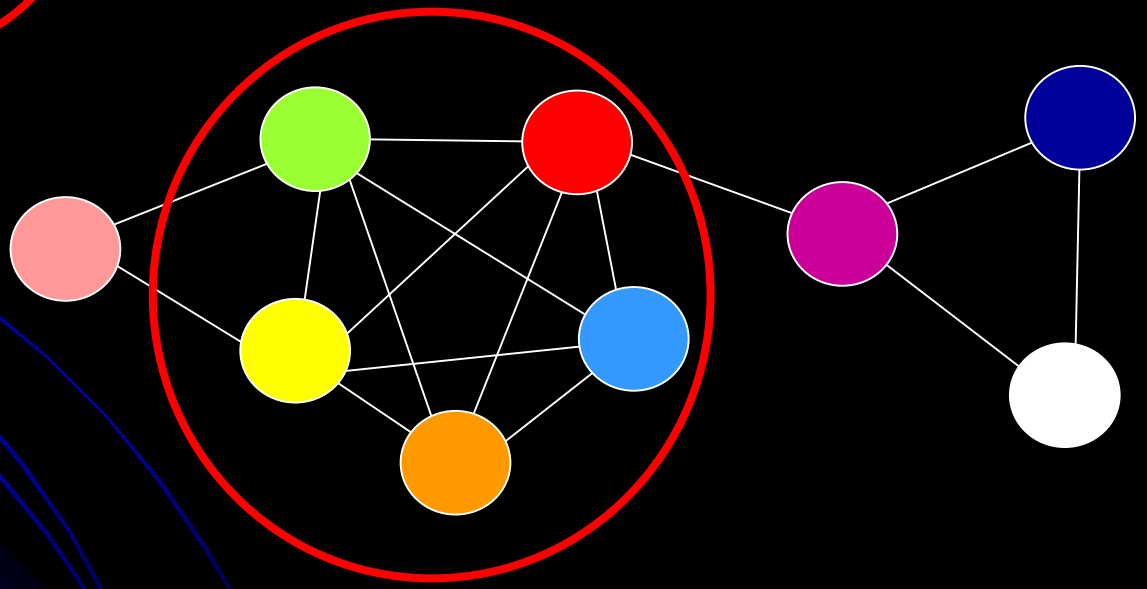
Overlap graph:
Nodes= functional groups
Edges= overlaps between them



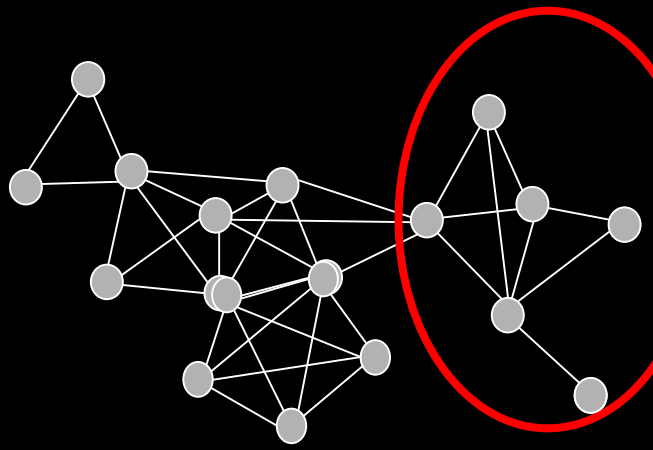
First line of attack



Overlap graph:
Nodes= functional groups
Edges= overlaps between them



First line of attack

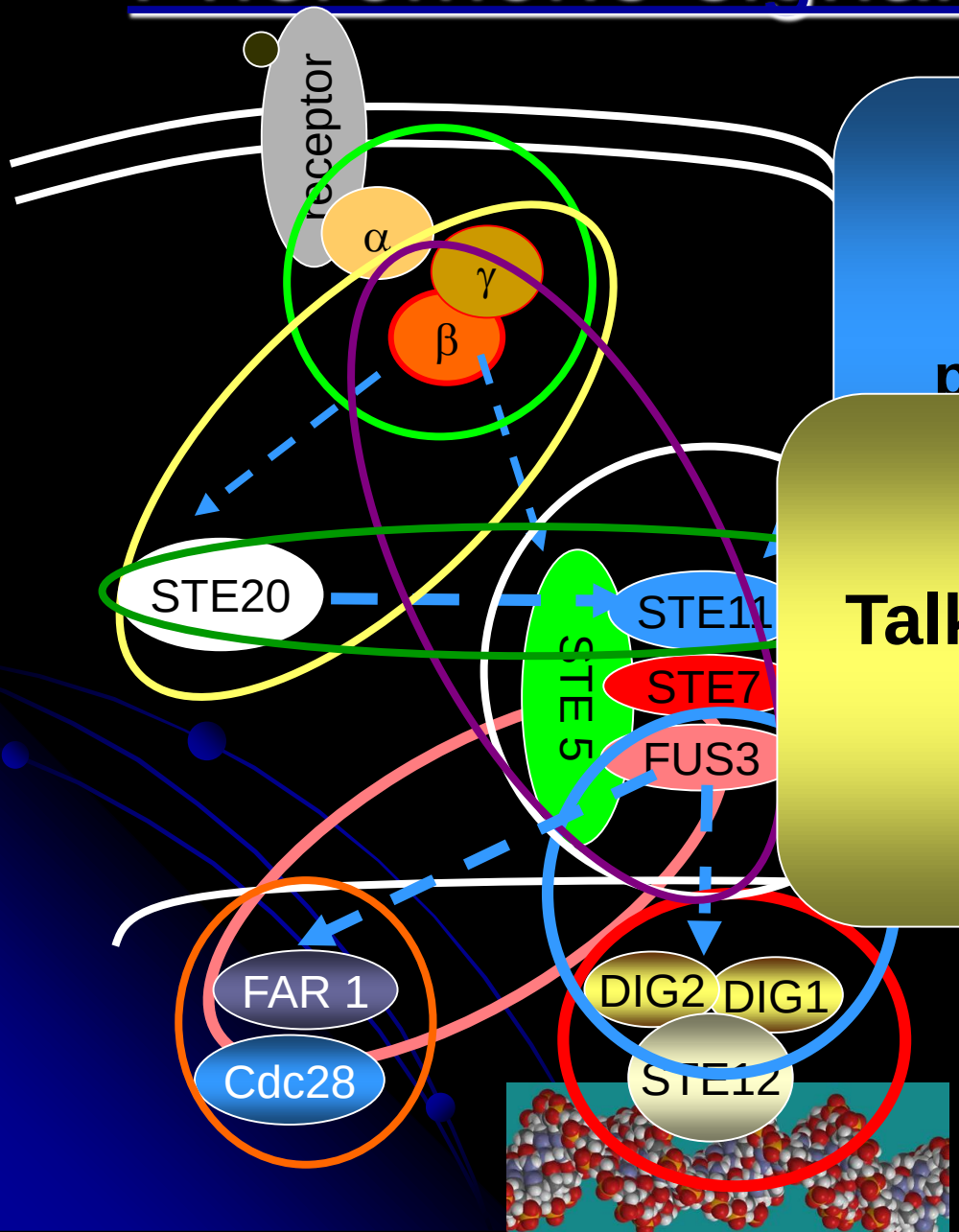


Overlap graph:
Nodes= functional groups
Edges= overlaps between them

Our contribution

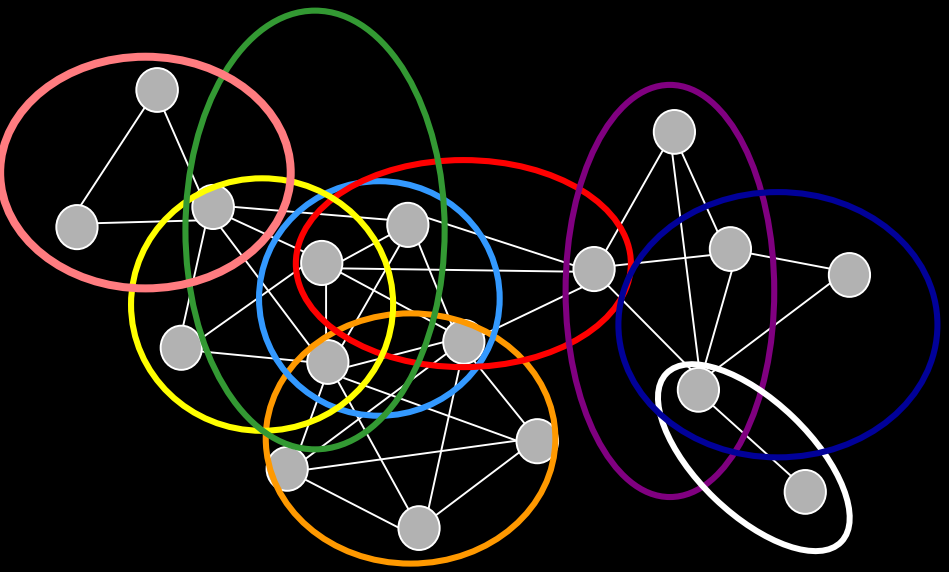
- Approach that elucidates the structure of the overlaps
- Application: If the functional module represents dynamically changing protein associations, the method can suggest temporal relations between these associations

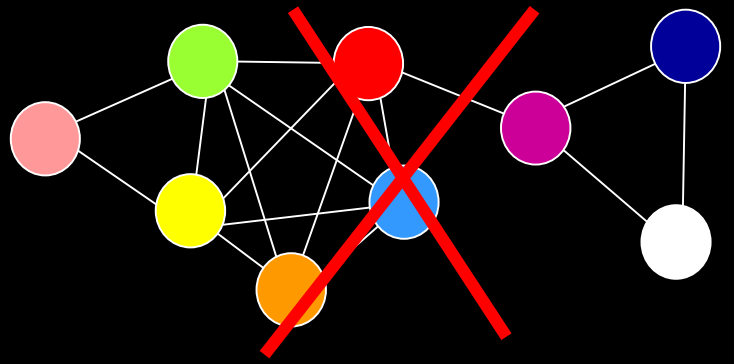
Pheromone signaling pathway



The information about overlaps be used to propose a partial order of

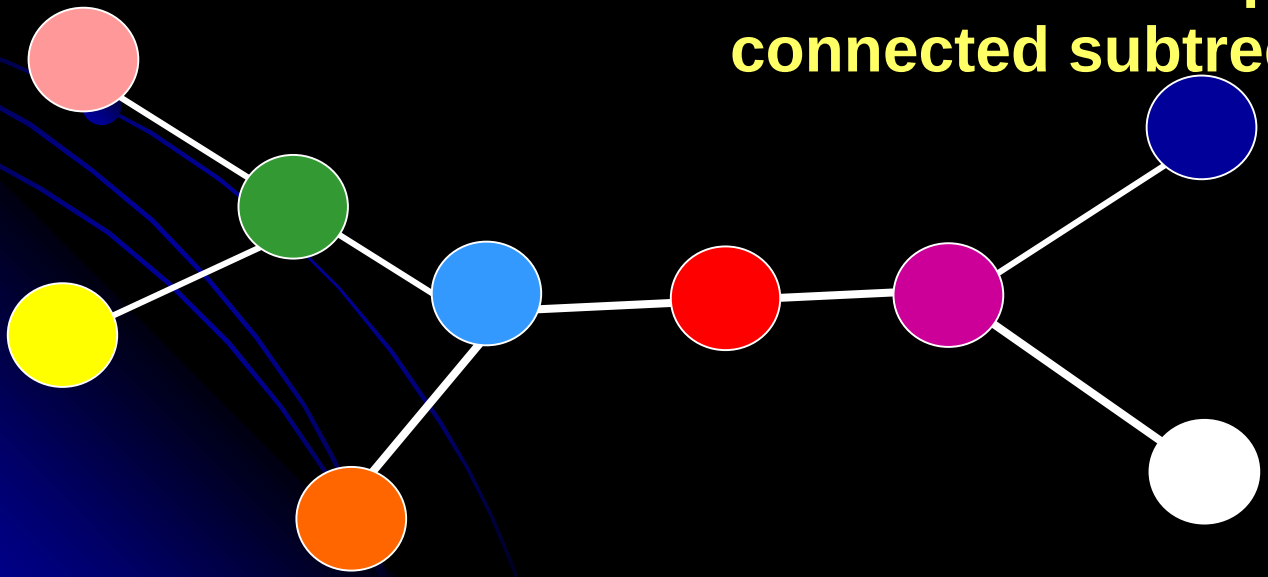
Talking partial order = talking trees



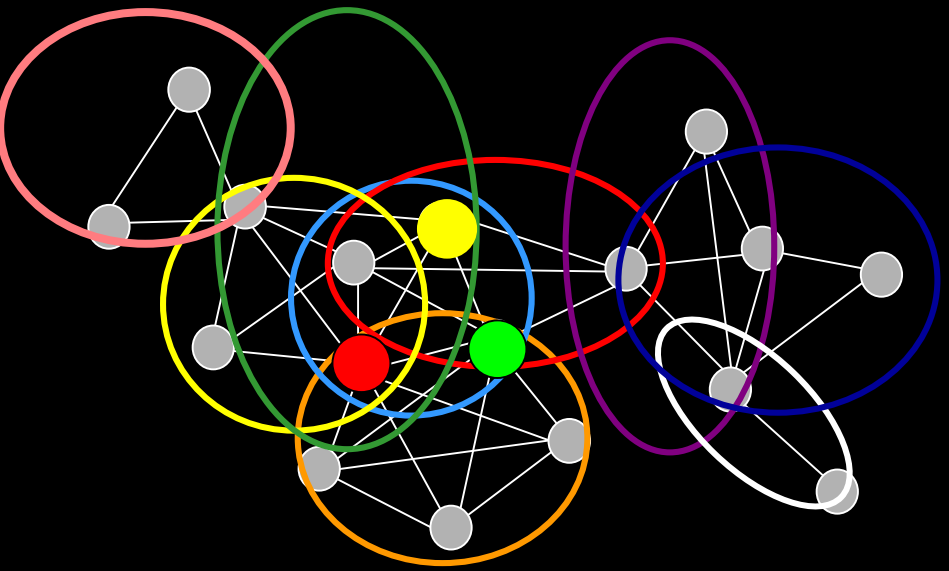


Clique tree

- Each tree node is a clique
- For every protein, the cliques that contain this protein form a connected subtree

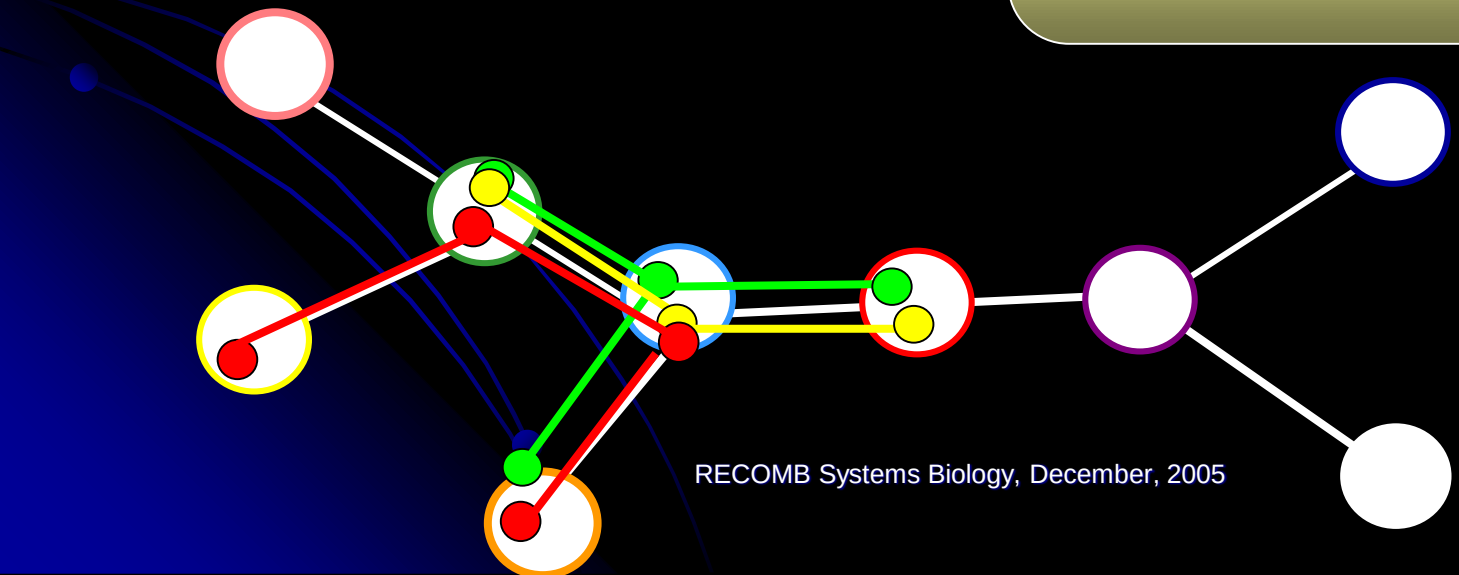


Key properties of a clique tree

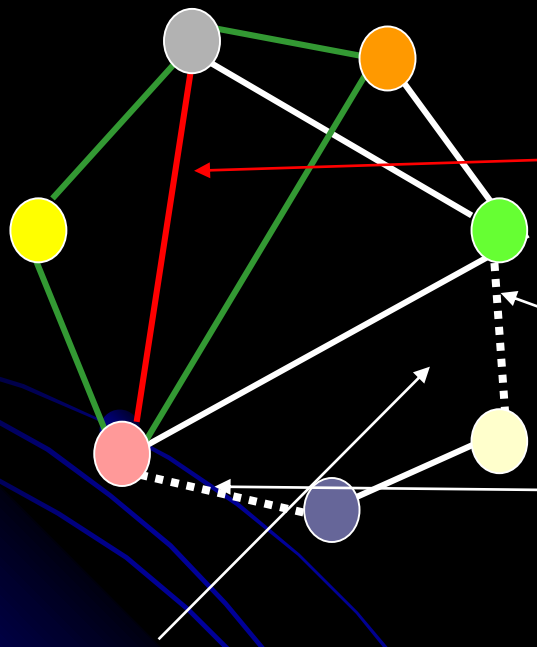


We can trace each protein as it enters/ leaves each complex (functional group)

Can such a tree always be constructed?



Clique trees can be constructed only for chordal graphs



Chord = an edge connecting two non-consecutive nodes of a cycle

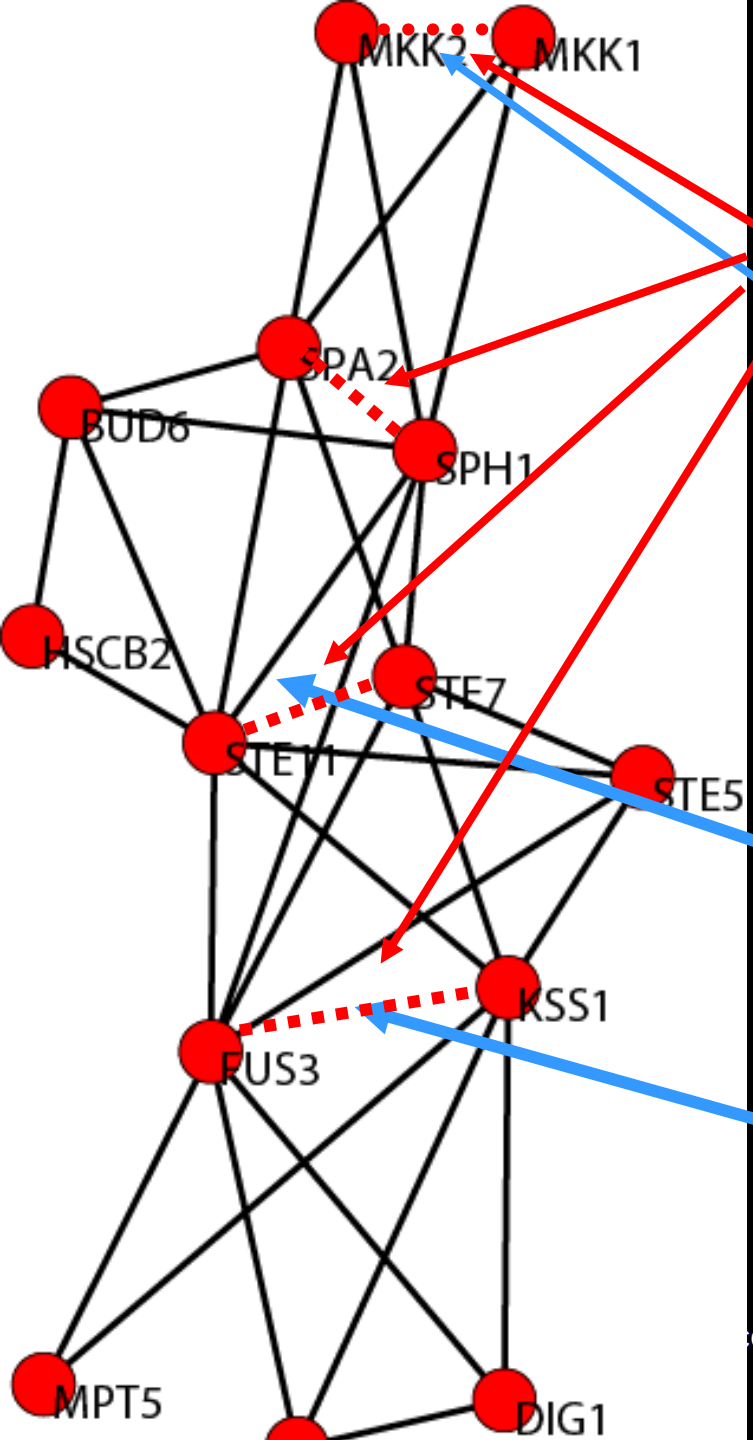
Chordal graph – every cycle of length at least four has a chord.

With these two edges the graph is **not** chordal

hole

Chordal PPI-networks???
Do we ask for too much?

- If we take the whole PPI network this IS asking too much.
- If we take a densely connected subgraph identified as a “functional module” by any of a number of approaches it is typically chordal or close to it!



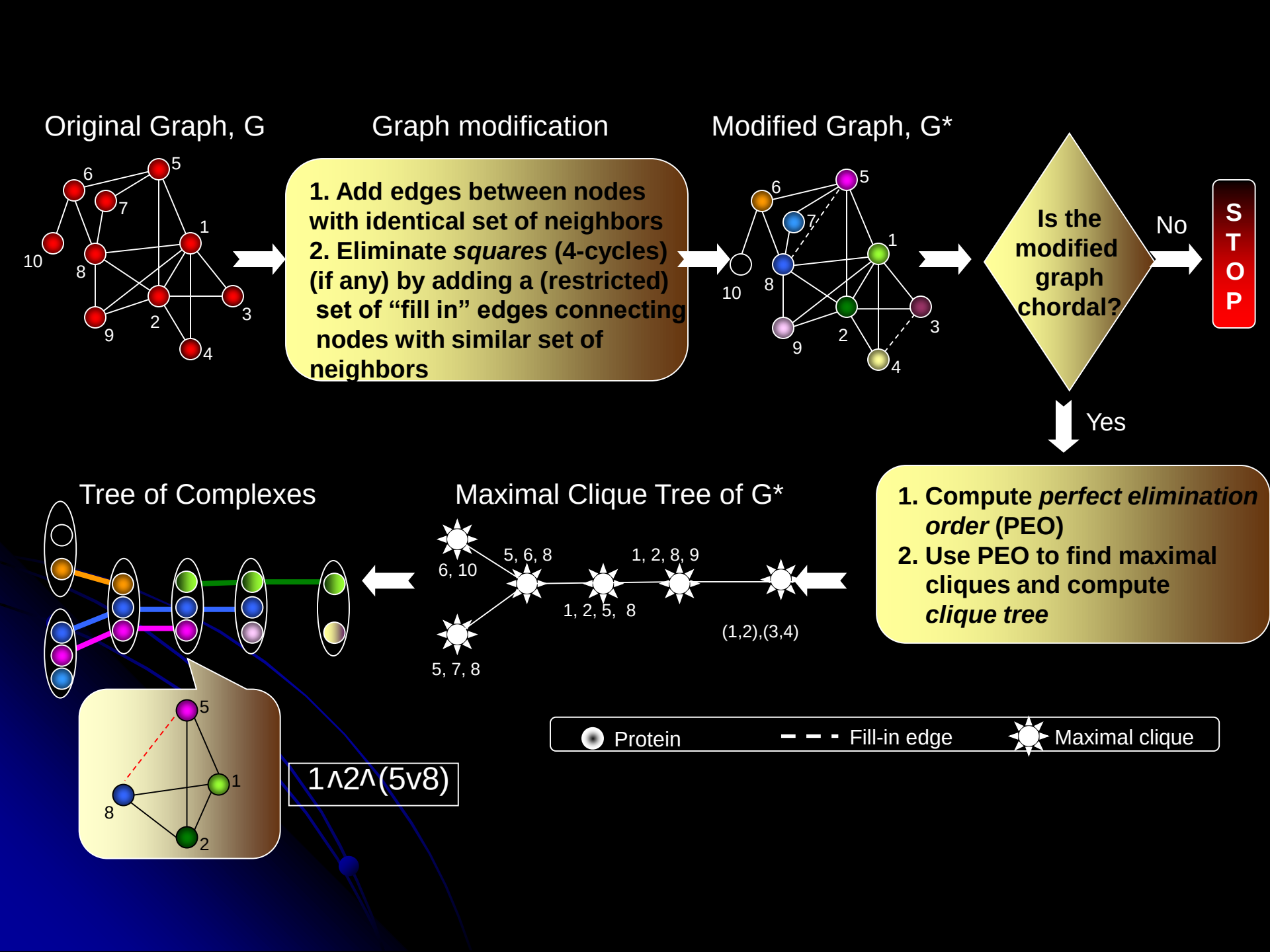
Add special "OR" edges

assembled by Spirin *et al*. 2004

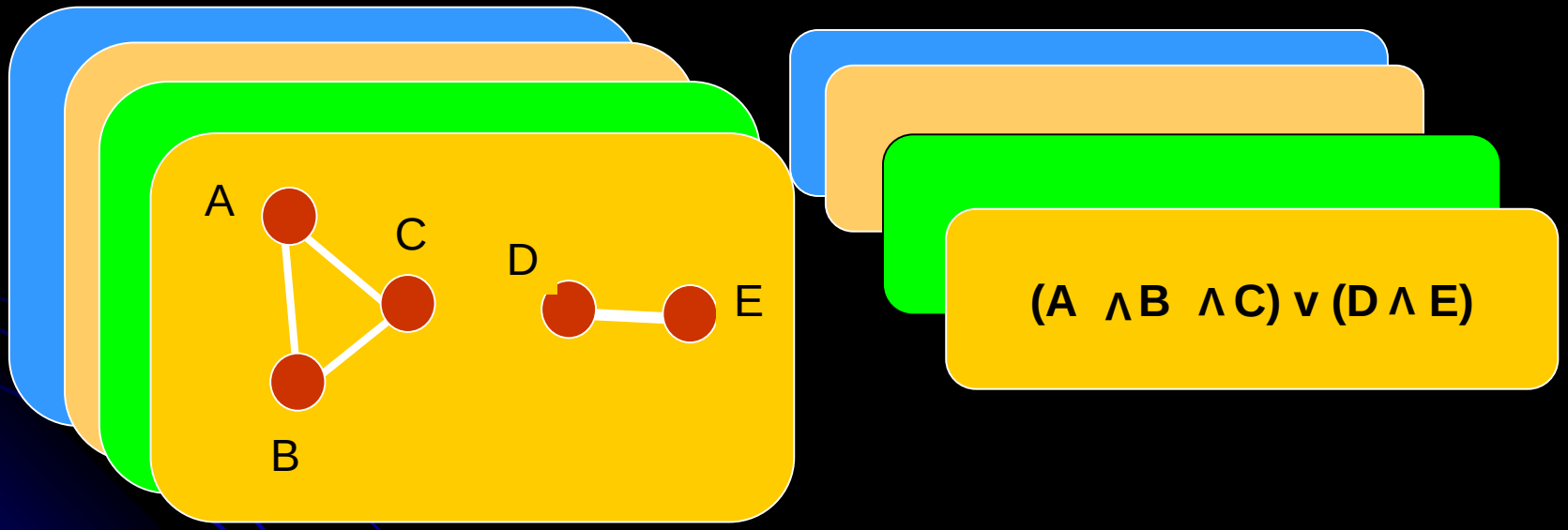
Square 1:
MKK1, MKK2 are experimentally confirmed to be redundant

Square 2:
STE11 and STE7 – missing interaction

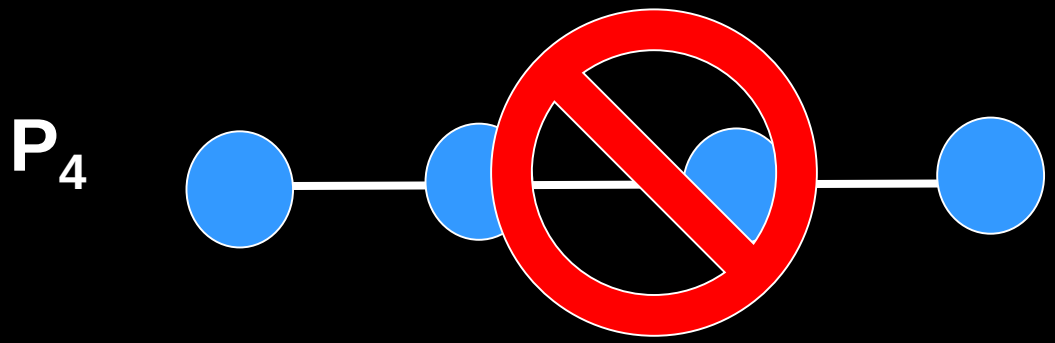
Square 3:
FUS3 and KSS1 – similar roles (replaceable but not redundant)


Original Graph, G
Graph modification
Modified Graph, G*
1. Add edges between nodes with identical set of neighbors
2. Eliminate squares (4-cycles) (if any) by adding a (restricted) set of "fill in" edges connecting nodes with similar set of neighbors
Is the modified graph chordal?
No
STOP
Yes
1. Compute perfect elimination order (PEO)
2. Use PEO to find maximal cliques and compute clique tree
Maximal Clique Tree of G*
6, 10
5, 6, 8
1, 2, 8, 9
1, 2, 5, 8
(1,2),(3,4)
5, 7, 8
Tree of Complexes
1∧2∧(5v8)
Protein
Fill-in edge
Maximal clique

Representing a functional group by a Boolean expression

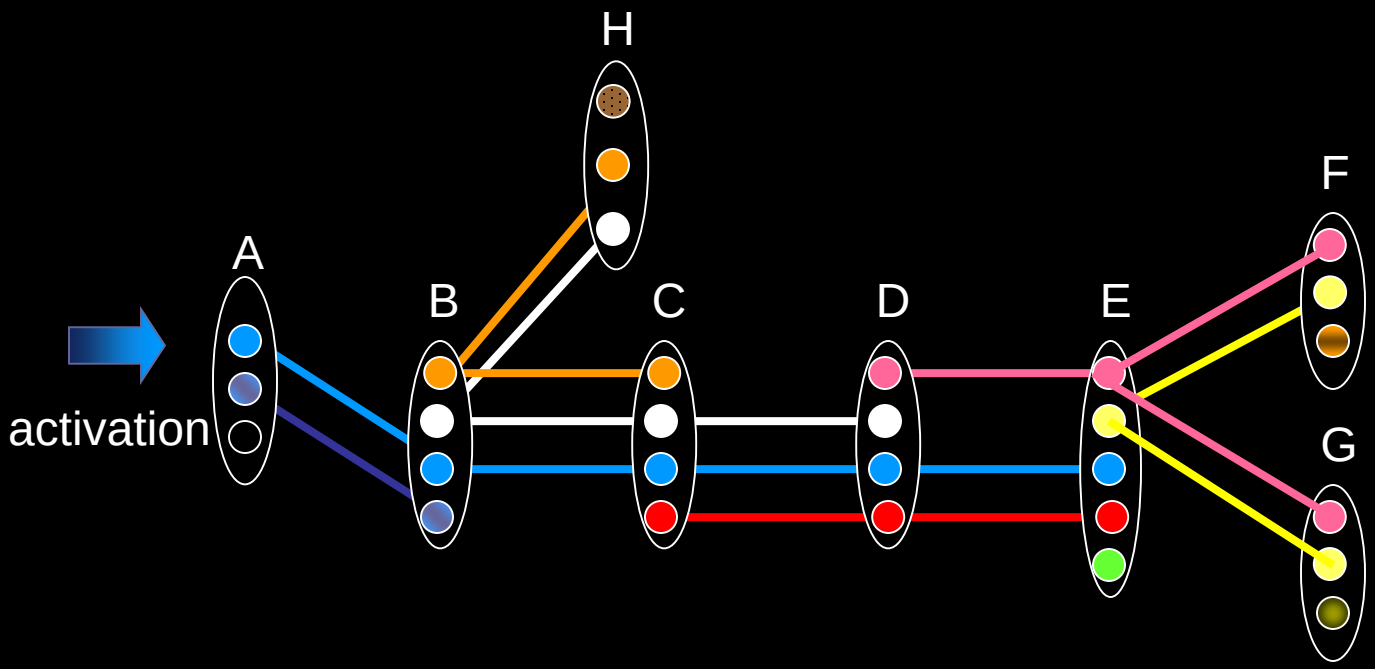


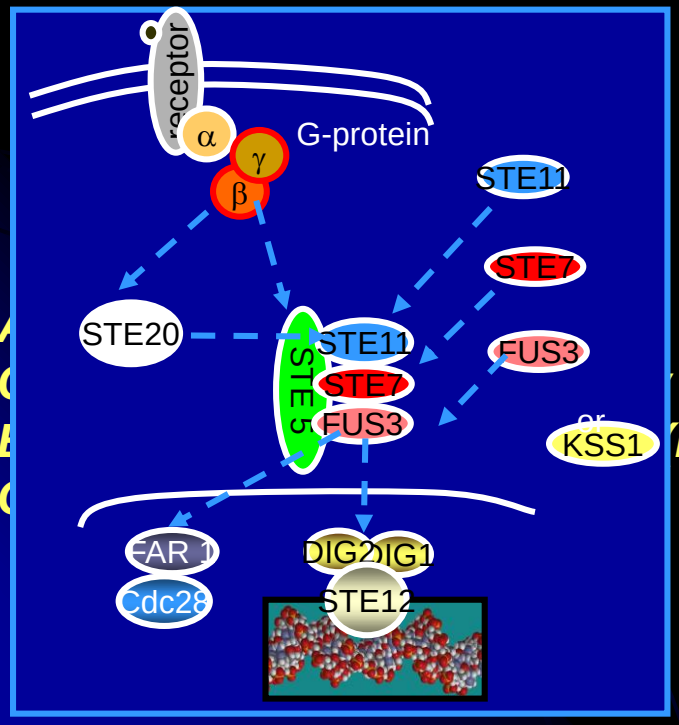
Not all graphs can be represented by Boolean expression



Cographs = graphs which can be ex

Connected components of a cograph are dense and "cliquish"





● = STE11 ● = FUS3 ○ = HSCB2
● = STE5 ● = KSS1 ● = BUD6
● = STE7 ● = DIG1^ DIG2 ● = MPT5

FUNCTIONAL GROUPS

B = BUD6 ^(SPH1 v SPA2) ^STE11
STE7) *D = SPH1^(STE11 v STE7) ^ FUS3*
(FUS3 v KSS1) *F = (FUS3 v KSS1) ^DIG1^ DIG2*
H = (MKK1 v MKK2)^(SPH1 v SPA2)

Summary

- We proposed a new method delineating functional groups and representing their overlaps
- Each functional group is represented as a Boolean expression
- If functional groups represent dynamically changing protein associations, the method can suggest a possible odder of these dynamic changes
- For static functional groups it provides compact tree representation of overlaps between such groups
- Can be used for predicting protein-protein interactions and putative associations and pathways
- To achieve our goal we used existing results from chordal graph theory and cograph theory but we also contributed new graph-theoretical results.

Future work

- So far we used methods developed by other groups to delineate functional modules and analyzed them or method. We are working on a new method which would work best with our technique.
- Considering possible ways of dealing with long cycles.
- Since fill-in process is not necessarily unique consider methods of exposing simultaneously possible variants.
- Add other information, e.g., co-expression in conjunction with our tree of complexes.

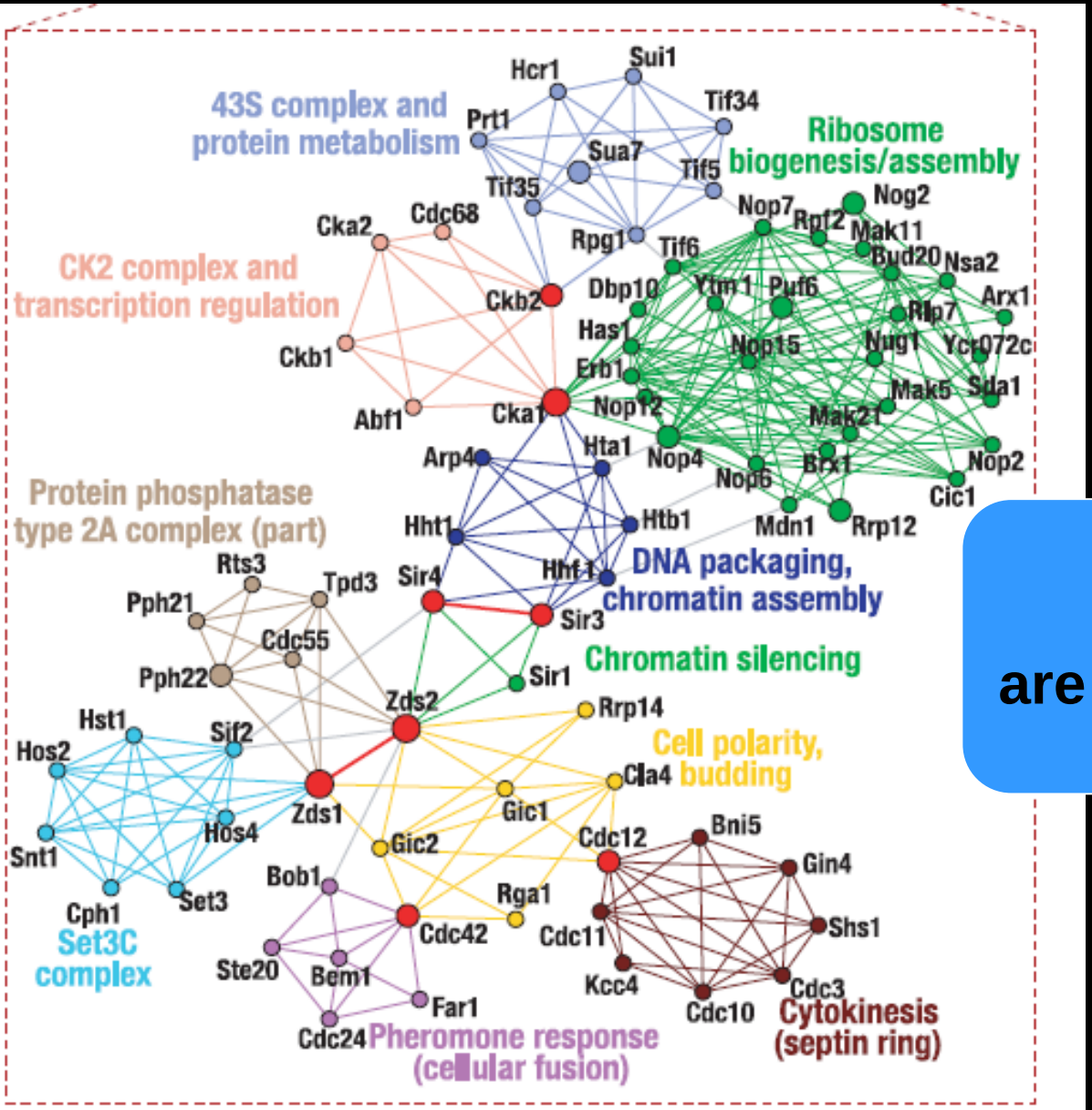
Acknowledgments

Elena Zotenko (UMD/NCBI)
Katia S. Guimaraes (Federal University of Pernambuco, Brazil / NCBI)
Raja Jothi (NCBI)

See also our poster

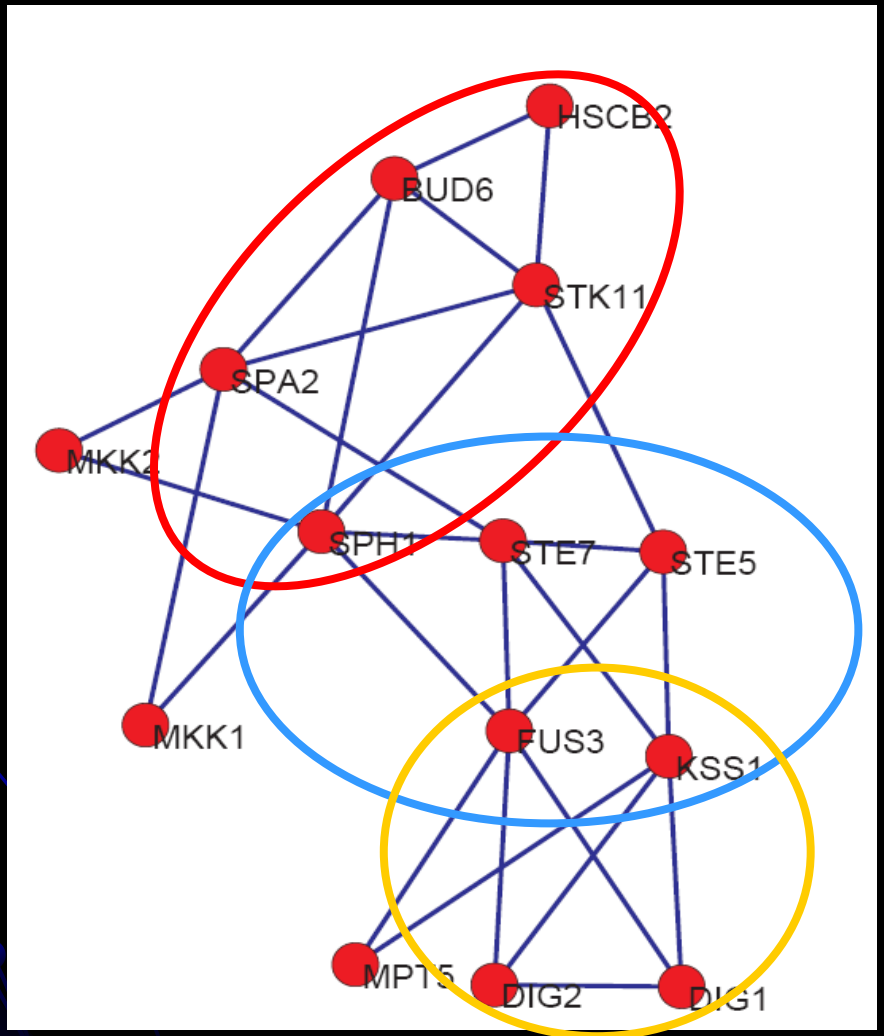
Applications

- "OR" edges – alternative/possible missing interactions. It is interesting to identify them and test which (if any) of the two possibilities holds
- Testing for consistency
- Hypothesis



Communities are typically chordal !

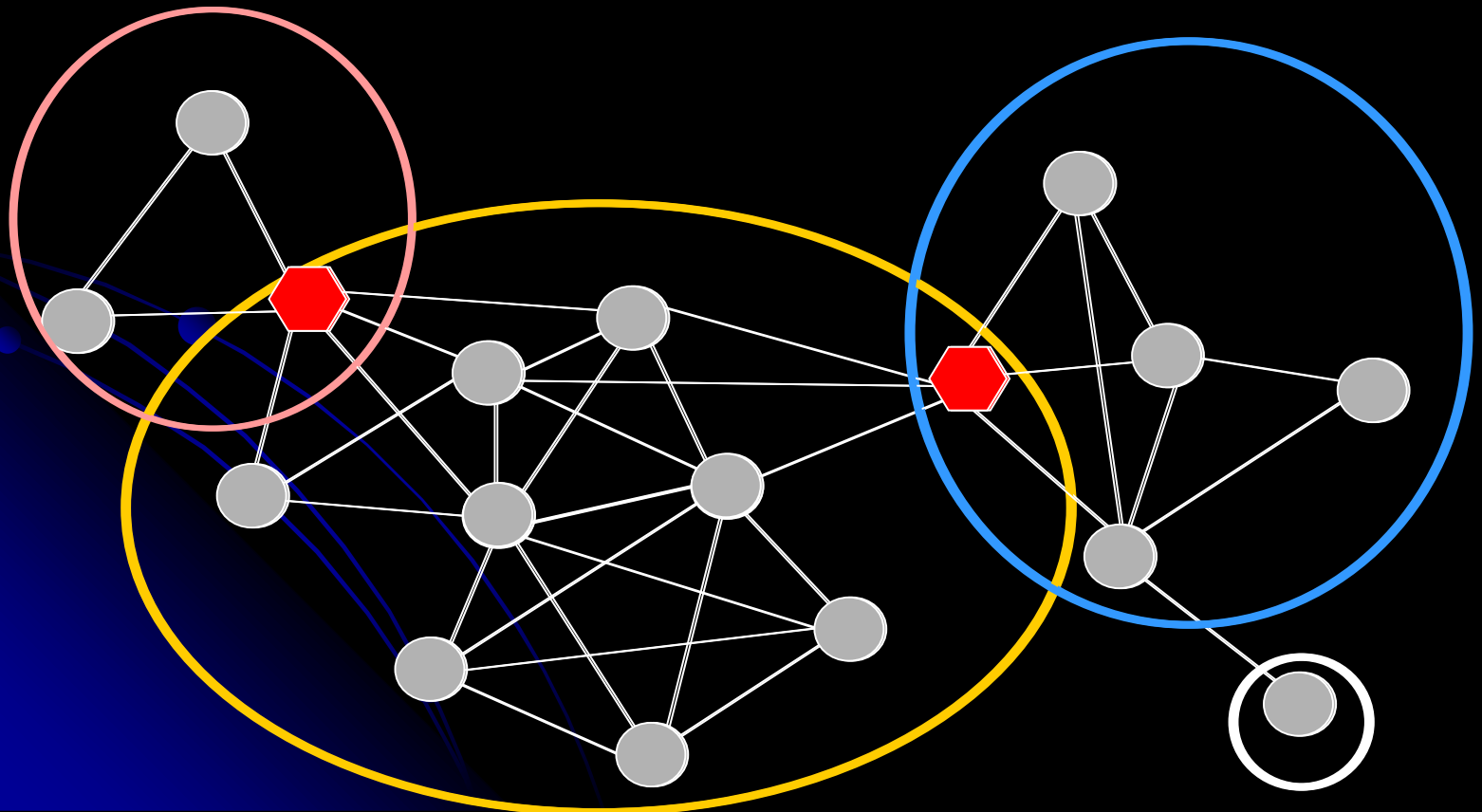
Community method applied to Pheromone data



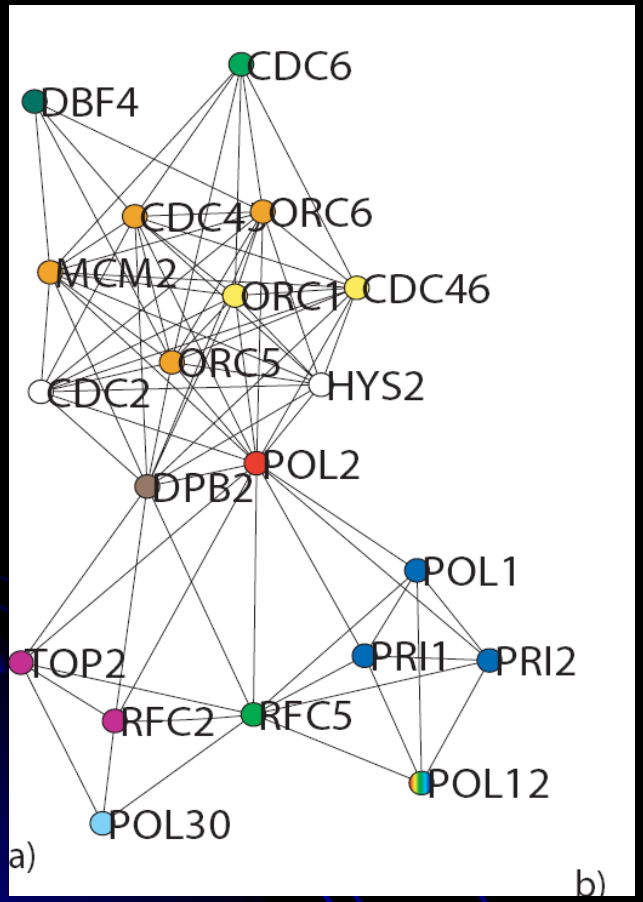
Palla et al.

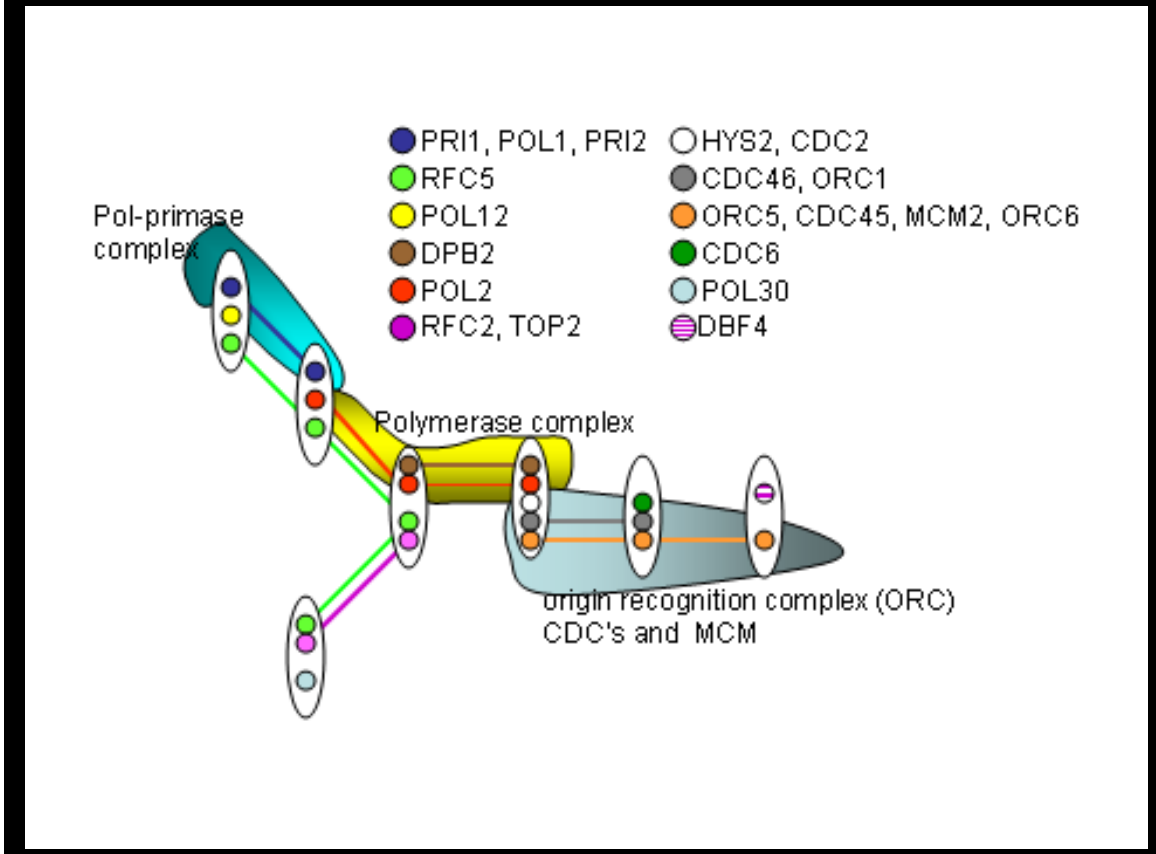
- Identify functional modules (allow them to overlap)
- For every node, record functional modules it belongs

(Palla *et al.* Nature, June 2005)



Transcription complex

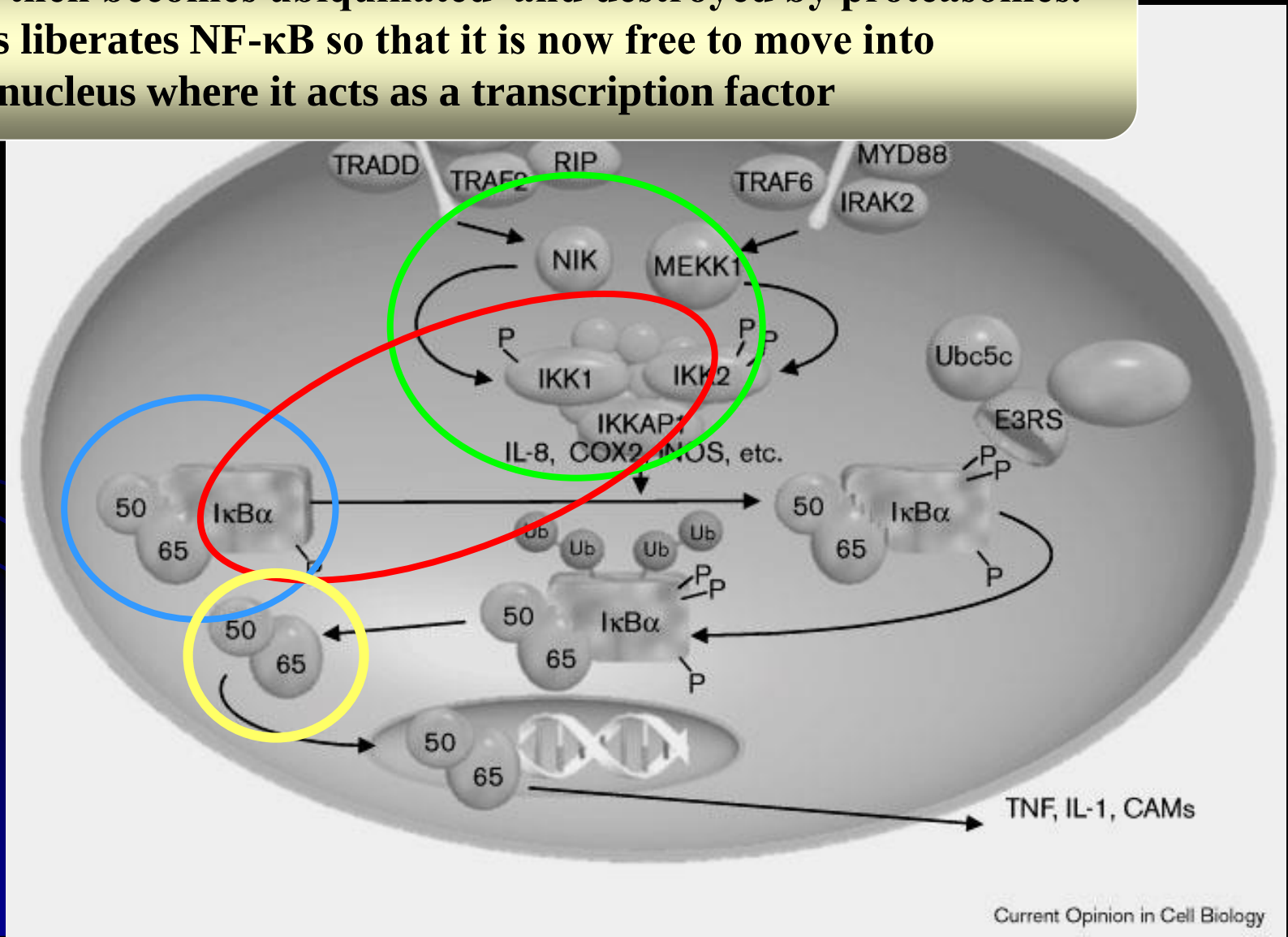




NF-κB resides in the cytosol bound to an inhibitor IκB.

Binding of ligand to the receptor triggers signaling cascade
In particular phosphorylation of IκB

IκB then becomes ubiquinated and destroyed by proteasomes. This liberates NF-κB so that it is now free to move into the nucleus where it acts as a transcription factor



Based on network assembled by:
Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, Cruciat C, Eberhard D, Gagneur J, Ghidelli S, *et al*.: A physical and functional map of the human TNF-alpha/NF-κB signal transduction pathway.
(all paths of length at most 2 from NIK to NF-κB are included)